**Engineering CRISPR guide RNA riboswitches for *in vivo* applications**

Roberto Galizi
r.galizi11@imperial.ac.uk

Alfonso Jaramillo, 1, 3
alfonso.Jaramillo@synth-bio.org

1Department of Life Sciences, Imperial College London, London, United Kingdom
2Warwick Integrative Synthetic Biology Centre (WISB) and School of Life Sciences, University of Warwick, CV4 7AL Coventry, United Kingdom
3ISSB, CNRS, Univ Evry, CEA, Université Paris-Saclay, 91025 Evry, France
4Institute for Integrative Systems Biology (I2SysBio), University of Valencia-CSIC, 46980 Paterna, Spain

CRISPR-based genome editing provides a simple and scalable toolbox for a variety of therapeutic and biotechnology applications. Whilst the fundamental properties of CRISPR proved easily transferable from the native prokaryotic hosts to eukaryotic and multicellular organisms, the tight control of the CRISPR-editing activity remains a major challenge. Here we summarise recent developments of CRISPR and riboswitch technologies and recommend novel functionalised synthetic-gRNA (sgRNA) designs to achieve inducible and spatiotemporal regulation of CRISPR-based genetic editors in response to cellular or extracellular stimuli. We believe that future advances of these tools will have major implications for both basic and applied research, spanning from fundamental genetic studies and synthetic biology to genetic editing and gene therapy.

---

**Introduction**

The microbial adaptive immune system CRISPR clustered regularly interspaced short palindromic repeats (CRISPR) offers a unique and highly effective toolbox for targeted genomic engineering, providing a versatile system to virtually target any nucleic acid sequence of choice with the sole engineering of a short fragment of RNA. In a seminal paper Jinek et al. proposed to harness the nuclease power of the type IIA CRISPR system of *Streptococcus pyogenes* utilizing a single CRISPR associated protein (SpCas9) and a single sgRNA, instead of the two RNAs and the large Cas protein complex found in the original host [1] (Figure 2a). Most of the CRISPR genome editing to date has been carried out with the type II nuclease CRISPR associated protein 9 from *Streptococcus pyogenes* (SpCas9). Redesigned versions were also generated by mutagenizing the key residues that contribute to endonuclease activity [2] and target specificity [3], by producing smaller protein variants [4] or assembling these into subunits to ease its delivery [5]. Furthermore, an increasing number of Cas9 orthologues and related effector proteins have been characterised from diverse bacterial species, some of which exhibit reduced molecular size and different target site specificities [6]. The recently characterised type II CRISPR-associated endonuclease Cpf1, also called ‘Cas12a’, for instance, shows distinguishing features from SpCas9. Cpf1 requires only a 42-nt CRISPR RNA (crRNA) to find its target instead of the ~100-nt gRNA for SpCas9, recognises a T-rich protospacer-adjacent motif (PAM) [7,8] that is 5′ instead of 3′ of the target site (precisely TTTV for Cpf1, NGG for SpCas9), has intrinsic RNase activity for the processing of multiple crRNAs from a single transcript [9,10] and makes staggered cuts (5′ overhang) whilst Cas9 makes blunt cuts in the genome [7]. CRISPR-associated endonuclease and related gRNAs can work heterologously in most species to produce mutagenesis, gene knockin/knockout and chromosome deletion/translocations. If the protein is engineered by knocking out each of its two nickase domains to generate the catalytically inert proteins ‘dCas9’ and ‘dCpf1’, the sgRNA:dCas9 (or sgRNA:dCpf1) complex will remain bound to its double-stranded DNA
(dsDNA) substrate [2]. As many other DNA-binding proteins, such as the popular zinc fingers, sgRNA:dCas9 or dCpf1 can work as a transcriptional repressor for CRISPR interfering (CRISPRi) without further engineering [11,12] or by adding transcriptional repression domains such as the Krüppel associated box (KRAB) [13,14]. Deactivated CRISPR-associated endonucleases can be engineered to function as transcriptional activators or epigenetic modifiers other than being utilized for high-resolution genome imaging [15] and DNA/RNA pull-down [16]. Hereafter we will refer to all these variants as CRISPR-editors. At least in theory, it is possible to reprogram CRISPR specificity to recognize any target region flanking a short PAM sequence by simply providing different gRNA-spacers. The PAM therefore represents the sole CRISPR-targetability constraint, which has been shown possible to broaden by mutating the PAM interacting domain of the protein [17,18]. These findings empowered over the recent years a vast range of applications from basic biology to biotechnology and medicine. Nonetheless, several challenges are still hampering the broad applicability, efficiency, specificity and safety of CRISPR-editing systems (Figure 1). Between these, the precise and robust spatiotemporal control of the nucleases activity which, at least for the majority of the current designs, relies solely upon the spatiotemporal specificity of the polymerase II promoter used to transcribe the endonucleases component (generally Cas9 or Cpf1) or it is strictly delivery-procedure dependent when protein or RNA are directly used for the editing [19]. Several strategies to engineer inducible CRISPR systems have been proposed over the recent years (some of these are summarised in Table 1). In the following we will focus on inducible CRISPR systems based on riboswitch regulation by giving an overview of the existing tools, propose new designs and highlight some of the current and foreseen applications.

![Figure 1 Applications, evaluation criteria and challenges of CRISPR-mediated editing.](image)

**Figure 1** Applications, evaluation criteria and challenges of CRISPR-mediated editing. (a) Venn diagram showing some of the main CRISPR applications in human health, environment, agriculture and biotech and evaluation criteria (outer circle) [94–96]. (b) CRISPR-editing challenges. Target site selection: target site and PAM specificity. Spatiotemporal activity: unexpected leaky expression or activity. Resistance to cleavage: pre-existing polymorphisms or de novo mutations generated by end-joining repairing of the target. Incidence of homology directed repair (HDR) versus error prone non-homologous end joining repair. Off-targets: promiscuous cleavage at genomic sequences that have significant similarity to the target. Delivery: particularly relevant for gene therapy applications where a safe and efficient delivery of the CRISPR-editing to specific tissues and cells is an essential requirement. Confinement and containment: physical, reproductive, ecological or molecular barriers to limit CRISPR-editing to the target organisms o species [93,97,98].
CRISPR-editors can be designed to function as: endonucleases, in the form of catalytically active Cas9/Cpf1; transcriptional activators (e.g. dCas9-VP64, dCas9-SAM, dCas9-Suntag, dCas9-VPR) generally referred as CRISPRa (CRISPR activation), transcriptional repressors (e.g. dCas9 only or dCas9-KRAB) also defined as CRISPRi (CRISPR interfering); DNA or histone modifiers, by fusing dCas9 or dCpf1 proteins to histone demethylases (LSD1), histone acetyltransferases (p300), DNA methyltransferases (DNMT3a) or DNA demethylases (TET1) [95,98,99].

**Riboregulation and riboswitches: current designs and applications**

*Natural and synthetic riboswitches*

Riboswitches are structured *cis*-regulatory segments of RNA, mostly found in bacterial messenger RNAs (mRNAs), capable of modulating the gene’s protein product in response to small molecules such as specific metabolites. These usually consist of an aptamer domain that binds with high specificity and affinity to a target ligand and transduces the binding signal into a gene expression output. Numerous RNA aptamers have been discovered, showcasing...
their intrinsic ability to tightly and selectively recognise specific molecules and to trigger ligand-dependent gene regulation in bacteria, plants and fungi [20]. Riboswitch engineering requires computational analysis and design of aptamers that recognise xenobiotic compounds as well as the development of new testing platforms. In natural riboswitches, the nucleotides within the aptamer domain responsible for ligand binding tend to be evolutionarily conserved and when mutated can result in changed ligand specificity, usually by altering the hydrogen-bonding pattern of the domain responsible for ligand recognition. A large number of riboswitch designs are indeed inspired by nature: first, natural purine (guanine and adenine) riboswitches [21,22], widely used to engineer orthogonal gene activation or repression counterparts to generate riboswitch-based controllers of cell behaviour and physiology in bacteria [23,24]; second, PreQ₁ riboswitches from a queuosine precursor, involved in the bacterial transfer RNA (tRNA) function [25,26]; third, flavin mononucleotide riboswitches from flavoenzyme, a natural product whose antibacterial activity involves the targeting of riboswitches [27]. Other classes, such as the cyclic diucnucleotide riboswitches (cddiG) were instead identified by bioinformatics analysis as for the highly conserved RNA structured element found upstream 'GEMM motif' (genes for environment, membranes and motility) [28].

The dynamic response and the high affinity for target ligands make riboswitch aptamers particularly suitable to function as recognition domain for the development of RNA-based fluorescent biosensors [29]. Here the riboswitch is fused to a transducer module that activates a signaling domain after ligand binding. Between these, Spinach, Spinach2 and Broccoli aptamers have been widely used to build ligand-responsive biosensors for live-cell imaging of target metabolites by producing a fluorescent output upon addition of a substrate [29-31]. Besides being widely used to engineer biochemical reporters (e.g. responsive to vitamin B₁₂ [32] and other enzyme quenches [33,34]), biosensors and for in vivo imaging of small molecules, riboswitches have also been engineered with xenobiotic molecules to control gene expression in a wide range of prokaryotic microorganisms (e.g. Escherichia coli, Acinetobacter baumannii, Bacillus subtilis, Mycobacterium smegmatis, Mycobacterium tuberculosis and other streptomycetes and cyanobacterial species [35-40]), yeasts [41-44], fungi [45], plants [46] and mammalian cells [47,48]. Riboswitches are also suitable for the engineering of more complex regulations through the implementation of biochemical logic gates and feedback systems [49-51]. RNA-based biosensors are also extensively used for the engineering of metabolic pathways to optimise the production of small molecules such as industrial and fuel compounds [52], or for the sensing of intracellular behaviours and environmental conditions [53].

**RNA-responsive riboswitches**

Synthetic trans-activating RNAs, generally called 'riboregulators', are one of the first examples of synthetic translational regulators described to function as ligands. When added to the 5' UTR (untranslated regions) these structures can be used to control gene translation by relaxing the structural blockade of hairpins designed to cis-repress ribosomal binding [54,55]. A re-engineered version named as 'toehold' [56] notably improved the range and the dynamic of post-transcriptional control for the development of RNA-nanodevices to evaluate complex logic circuits in living cells [57]. Longer mRNA transcripts can also function as translational triggers [56] providing the means to build low-cost diagnostic platforms based on toehold designs [58].

**Riboswitch designs and validation**

Rational or semi-rational designs can be performed by decoupling expression platforms from the activity of the natural aptamer domain followed by the integration of foreign aptamers that are responsive to different ligands [59]. A major challenge faced during the aptamer screening and selection phase is to address the requirement of maintaining functional gene regulation in vivo [60]. For instance, the most common methods to generate riboswitches, such as the in vitro systematic evolution of ligands by exponential enrichment (SELEX), can generate aptamers and expression platforms that do not always function in vivo [61,62] and therefore require downstream screening steps for a comprehensive characterization of their outputs [63].

Computational modeling offers a valuable option to generate functional riboswitches for in vivo applications by using the folding energy of the riboswitch combined with the free energy change resulting from ribosome binding to the RBS [35] or by analysing co-transcriptional RNA–ligand interaction dynamics [64,65,66]. A number of expression platforms have been described over the last few years and classifiable as: first, co-transcriptional regulation, where the ligand-binding attenuates transcription by forming a terminator hairpin (e.g. PreQ₁ riboswitches) [25,26,67] or interfering with rho dependent transcription termination (e.g. the E. coli thiM riboswitch and the Salmonella corA) [68,69]; second, inhibition of translation, with ligand binding forming an alternative structure that occludes the ribosome-binding site (RBS) and prevents initiation of translation (e.g. thic riboswitch) [70]; third, mRNA degradation, caused by ligand-induced ribozyme activity (e.g. the Bacillus subtilis glyS RNA) [71]. Post-transcriptional RNA splicing regulation, where the ligand binding leads to changes in splice sites to form alternate mature mRNA by addition of open reading frames (ORF), inclusion of premature stop codon (e.g. thiamin pyrophosphate (TPP) riboswitches) [72,73] or via excision of the polyadenosine (poly-A) tail followed by mRNA degradation (e.g. flowering plants THIC genes) [74,75]. Interestingly, TPP riboswitches are the sole type of natural riboswitches being described in plants, marine protists and fungi [42]. There are examples of artificial expression platforms shown to function in eukaryotes such as: first, ribosome blocking, where the aptamer–ligand interaction inhibits the ribosome scanning from the 5' cap to the translation start site preventing translation in yeast (it is yet to be confirmed whether the mammalian ribosomes may or not be able to scan through structured RNAs positioned upstream the start codon [76]); second, ligand-responsive ribozymes (also called ‘aptazymes’) to promote cleavage of 5' UTRs (untranslated regions), where the loss of the 5' cap prevents ribosome initiation, or 3' UTRs followed by loss of the poly-A tail and degradation of the mRNA transcript [47,48,77]; third, aptazymes associated to micro RNA (miRNA) precursors [78,79] or orthogonal tRNAs for unnatural amino acid incorporation and studies of ribosome function (e.g. T-box riboswitches) [37].

Riboswitches and toehold riboregulators have been engineered for a wide number of applications from basic genetic studies to the production of synthetic compounds and diagnostic tests [53,80-82]. Beyond addressing the technical limitations,
one of the major and most exciting challenges is the development of novel approaches that would allow to broaden the use of riboswitches to eukaryotic systems, currently hindered by the cell-specific properties of aptamers or ligands, and the significant differences between the transcription and translation machinery employed by prokaryotic and eukaryotic systems.

In the following, we will describe how novel riboswitch designs can be combined with the pre-existing CRISPR-editing components to engineer specific and sensitive molecular tools to achieve targeted and inducible genetic editing in a wide range of prokaryotic and eukaryotic hosts. We believe that these technologies could be potentially transferrable to any CRISPR-based application.

**CRISPR gRNA riboswitches: new designs and future applications**

Riboswitches have been extensively engineered to control gene expression by incorporating aptamers that, after ligand binding, produce a conformational change to trigger the expression of a downstream gene. A similar strategy can be applied to sense CRISPR-sgRNAs (Figure 2a) by grafting an aptamer domain into an allosteric structure that hybridises an obliterates the sgRNA binding to the target. The interaction with the ligand stabilises the aptamer to create a conformational change where the guide sequence is fully exposed. This has already been tested successfully in eukaryotes using Cas9 [83] or Cpf1 [14]. We have also adapted theophylline sensing designs previously developed in our lab [77,84] to generate sgRNAs in *E. coli* (unpublished) (Figure 2b). Aptazymes developed to create UTR-linked riboswitches (described above) can be adapted to control sgRNA activity by adding the corresponding sequence to its 5′-end to repress the guide sequence in the unbound conformation and reconstitute the active conformation after ligand binding (Figure 2c) [85**,86**]. Potentially any ligand-responsive riboswitch could be adapted for sgRNA sensing. For example, light or temperature sensing sgRNA could be generated by replacing the aptamers in the designs represented in Figure 2b and 2c with an aptamer that interacts with a photo-switchable molecule [87] or using temperature riboswitches to regulate the initiation of translation or the transcription termination [88]. The main limitation of CRISPR-based riboswitches is represented by the nature and properties of the ligand–aptamer interaction. If RNA molecules are adapted to function as a trigger to sense sgRNA activity, these molecular tools could be utilized to regulate CRISPR-editing activity in response to endogenous or exogenous RNAs and, at least in theory, could be used for any application and organism where specific spatiotemporal or conditional regulation may be required.

*Figure 2* CRISPR gRNA riboswitch designs. (a) Diagram showing the interruption of transcription elongation when the dCas9 (or dCpf1, details not shown here) endonuclease binds to the PAM sequence, unwinds the dsDNA and the guide sequence region c of the gRNA binds to its target in the non-template strand or generates a double-strand break when the active Cas9 (or Cpf1) is used. (b) The gRNA is designed with an extra sequence at the 5′-end that hybridises with the guide sequence, which prevents the binding of the gRNA-endonuclease complex to its target. Addition of the theophylline ligand during gRNA transcription stabilises an open conformation with free guide sequence, which will repress transcription after binding to the inactivated protein or cleavage with the active nuclease. (c) Design of a gRNA with a non-functional ribozyme on its 5′-end hybridising with the guide sequence. Binding of the ligand stabilises the ribozyme, which self-cleaves and the gRNA with free guide sequence will complex to the protein for repression or cleavage. These designs could be also adapted to dCas9 or dCpf1 linked to activators (e.g. VP64, SAM, SunTag, VPR), repressors for CRISPR interfering (e.g. KRAB), DNA or histone modifiers (e.g. histone demethylase LSD1, histone acetyltransferase p300, DNA methyltransferase DNMT3a, DNA demethylase TET1).
With this aim, we designed cis-repressed switching-sgRNA, termed 'interacting guide RNA' (igRNA) that change conformation and switch to their active forms upon the binding of trigger-RNAs (trRNA) in the form of sRNA or mRNA. The first approach is based on previously described synthetic riboregulator properties [89], where the trRNA-igRNA pair is analogous to a trans activating riboregulator-5'UTR pair. Here the trRNA strand directly hybridises with the guide counterpart hampering the flexibility of guide and trigger sequence design (Figure 3a). To overcome this limitation the stem loop can be redesigned to flank the clamp sequence with two variable regions that hybridise with the trigger sequence (Figure 3b). The strategies in Figure 2b and c could also be adapted by replacing the ligand with a trRNA to stabilise the sgRNA stem to its open conformation (Figure 3c and d). Leveraging the intrinsic multiplex capacity of the Cpf1 platform, the designs proposed in Figure 3 could also be adapted to Cpf1 gRNAs to streamline the engineering of complex genetic circuits for gene perturbation studies [90–92] and for a number of other applications where inducible and multiplexed CRISPR-editing activity may be required (Figure 1). We are currently applying our RNA-triggered igRNA designed to sense the cell state in E. coli and in mosquito vectors of infectious diseases to activate CRISPR-Casendonucleases in response to tissue specific endogenous transcripts.

Figure 3 ENDogenous or exogenous RNA could be sensed by a switchable ‘interacting guide RNA’ (igRNA). (a) Design of a gRNA with upstream sequence hybridising against the guide sequence. When a trigger RNA (trRNA) such as mRNA, sRNA, gRNA etc. is present during gRNA transcription, the trigger sequence binds to the gRNA stabilising a conformation where the guide sequence is freed. The complex binds to the dCas9 (or dCpf1) to repress transcription of a gene containing the target sequence (that is also complementary to the trigger in this case) or to generate double-strand break if the active endonucleases are used. (b) Modification of the previous design, where part of the guide sequence is initially hybridised, but the regions complementary to the trRNA are only in the exposed 5'-end and loop, and not in the region hybridising to the target sequence. This allows the design of universal target sequences with no trRNA or gRNA-target constraints. (c) The strategy in Figure 2b, where the ligand (here the trRNA) stabilises a conformation in which the gRNA sequence is no longer hybridised. (d) The strategy in Figure 2c, where a destabilised hammerhead ribozyme gets reconstituted after binding to the trRNA. As for the configurations represented in Figure 2, also these designs could be adapted to any other type of Cas9 or Cpf1 CRISPR-editor that may function as inducible endonuclease, transcriptional activator or repressors, DNA or histone modifier.

Conclusions

Non-coding RNAs have been engineered for a wide range of applications and it has been shown recently that similar techniques can be used to engineer gRNAs able to interface with the cellular environment by sensing small-molecules, temperature and light. We have described gRNAs with allosteric properties that allow for signal transduction, cascading, computation, as well as the monitoring and actuation of the transcriptome. The CRISPR-gRNA strategies described here could also be combined among themselves or with other regulatory systems (Table 1) to tackle complex roadblocks that are still hindering CRISPR-editing technologies (Figure 1). We envision that igRNA
riboswitches have great potentials for broadening the variety of applications beyond the development of components for standard biotechnology purposes. CRISPR-based riboswitches represent a versatile in vivo post-transcriptional regulation toolbox for the engineering of customizable and highly specific CRISPR-editors that can be designed for targeted genome editing, transcriptional activation/repression or epigenome modification and imaging, exclusively within specific cells, tissues or species where the triggering molecular component is present. The CRISPR-igRNA designs overviewed in Figure 3 could be used to achieve spatiotemporal control of eukaryotic genome editing in response to RNA transcripts that may be specific for: first, pathogenic conditions, such as invading viral RNAs, mRNA transcripts carrying photogenic mutations or produced upon cellular stress conditions; etc.; second, tissue/cell, developmental stage or condition, for applications spanning from gene therapy all the way to the genetic control of invasive species such as malaria transmitting mosquitoes. Here the igRNA could be designed to generate more effective CRISPR-based gene drives by restricting the CRISPR-nuclease activity to the germline [93], generate self-contained gene drive systems active only in laboratory strains (e.g. by using synthetic mRNA triggers not present in wild-type species) or even to engineer suicidal transgenes that respond to specific adverse gene mutations (such as the ones causing insecticide resistance in the mosquitoes).

Whilst we believe that ad hoc computational tools and novel in vivo based selection methods will be required, we envision that the development of CRISPR-gRNA based riboswitches designs described here will be extremely valuable to tackle most of the challenges that are still impeding the deployment of CRISPR-mediated genome editing and engineering.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- ** of outstanding interest

Acknowledgements

This work was supported by grants from the European Commission FP7-ICT-FET-610730 (EVOPROG), FP7-KBBE-613745 (PROMYS), H2020 Marie Sklodowska-Curie 642738 (MetaRNA), EPSRC-BBSRC BB/M017982/1 (WISB center), BBSRC grant EVO-ENGINE BB/P020615/1 and the startup allocation from the School of Life Sciences.

References


42. P.D. Donovan, L.M. Holland, L. Lombardi, A.V. Counihan, D.G. Hijmans, K.H. Wolfe and G. Butler, TPP riboswitch−dependent regulation of an ancient thiamin transporter in *Candida*, *PLoS Genet* **14**, 2018. The thiamin pyrophosphate (TPP) riboswitches are widespread in the budding yeasts (*Saccharomyces cerevisiae* and common in homologs of *DUR31*, including *Candida parapsilosis* were the authors demonstrate its molecular function.


Please add these two citations as highlighted in the text: ••"https://doi.org/10.1073/pnas.1203831109" followed by description "Automated design and validation of riboregulatory devices to create complex logic circuits in living cells" and ••"https://doi.org/10.1093/nar/gkx698" followed by description "Implementation of portable sRNA-regulated networks for the engineering of complex cellular programs and biological computations"").


Computational design of small transcription activating RNAs for versatile and dynamic gene regulation, Nat Commun 8, 2017.


First demonstration of transcriptional pausing at translation start sites described within the E. coli thiamin pyrophosphate (TPP) thiC riboswitch that acts as a checkpoint for thiC gene regulation.

Mechanism of mRNA destabilization by the glmS ribozyme, Genes Dev 21, 2007, 3356-3368.


Riboswitch control of gene expression in plants by splicing and alternative 3’ end processing of mRNAs, Plant Cell 19, 2007, 3437-3450.

Control of mammalian translation by mRNA structure near caps, RNA 12, 2006, 851-861.


Control of bacterial translation by mRNA structure near caps, J Appl Genet 57, 2016, 531-541.

Graphical abstract

CRISPR-editors can be reprogrammed for cleavage of nucleic acids (C), transcriptional activation (A) or repression (R) and epigenetic modification (M). Examples of gRNA designs (multiplexed and cis-repressed interacting gRNA). Chemical or biological (proteins or RNAs) ligands for regulation. Computational design and testing of gRNA-ligand interactions.
Highlights

- Riboswitches and CRISPR-editing challenges, current and future applications.
- Riboswitches can be applied to sense CRISPR-sgRNAs.
- CRISPR gRNA riboswitches can function in all kingdoms of life.
- Combined in vivo screening and computational analysis to improve riboswitch regulation.

Queries and Answers

Query: The author names have been tagged as given names and surnames (surnames are highlighted in teal color). Please confirm if they have been identified correctly.

Answer: Yes